USEWOD - Usage Analysis and the Web of Data

6th International Workshop on Usage Analysis and the Web of Data: The diverse ecosystem of Web of Data access mechanisms

Co-located with WWW 2016; 11th April 2016, Montreal, Canada (Venue and travel information)

PROCEEDINGS

PROGRAMME

9:00 - 9:05: Workshop opening

9:05 - 10:00: Keynote talk: **Beyond Well-designed SPARQL.** Mark Kaminski, University of Oxford.

Abstract: SPARQL is the standard query language for RDF data. The distinctive feature of SPARQL is the OPTIONAL operator, which allows for partial answers when complete answers are not available due to lack of information. However, optional matching is computationally expensive - query answering is PSPACE-complete. The well-designed fragment of SPARQL achieves much better computational properties by restricting the use of optional matching - query answering becomes coNP-complete. However, well-designed SPARQL captures far from all real-life queries - in fact, only about half of the queries over DBpedia that use OPTIONAL are well-designed. In the present paper, we study queries outside of well-designed SPARQL. We introduce the class of weakly well-designed queries that subsumes well-designed queries and includes most common meaningful non-well-designed queries: our analysis shows that the new fragment captures about 99% of DBpedia queries with OPTIONAL. At the same time, query answering for weakly well-designed SPARQL remains coNP-complete, and our fragment is in a certain sense maximal for this complexity. We show that the fragment's expressive power is strictly in-between well-designed and full SPARQL. Finally, we provide an intuitive normal form for weakly well-designed queries and study the complexity of containment and equivalence.

10:00 - 10:30: Short paper presentation: **Embedded Metadata and the Digital Lifecycle of Images: Methodological Challenges.** Nathalie Casemajor. (paper)

10:30 - 11:00: Coffee break

11:00 - 11:30: Research paper presentation: **Modeling Complex Clickstream Data by Stochastic Models: Theory and Methods.** Choudur Lakshminarayan, Ram Kosuru and Meichun Hsu. (paper)

11:30 - 12:00: Research paper presentation: **What public transit API logs tell us about travel flows.** Pieter Colpaert, Alvin Chua, Ruben Verborgh, Erik Mannens, Rik Van de Walle and Andrew Vande Moere. (paper)

12:00 - end: Breakout discussion: **The Web of Data ecosystem: What are the implications for usage analysis?**

ABOUT THE KEYNOTE SPEAKER

Mark Kaminski is a Senior Research Assistant at the Information Systems Group, Department of Computer Science, University of Oxford. His most recent research interests include efficient query answering and analytical query processing over RDF data, ontology-mediated query answering, as well as other ontology reasoning tasks such as module extraction and classification.

More generally, he is interested in all topics relating to modal, temporal and description logics, knowledge representation and reasoning, deductive databases, and semantic technologies.

He obtained his doctoral degree from Saarland University in 2012 working on tableau-based decision procedures for expressive modal and description logics. Besides, his previous research includes topics such as SAT-based reasoning or deductive completeness in higher-order logic.

THE 2016 USEWOD DATASET

As in previous years, a standard research dataset of usage data from well-recognized Web-of-Data datasets will be published to promote reproducible research on the workshop themes. **A particular highlight of this year's dataset will be overlapping usage data from the official DBpedia servers as well as the Linked Data Fragments interface to DBpedia and Wikidata.** This dataset allows researchers to study alternative Web of Data usage mechanisms in an unprecedented way and could therefore become a unique resource of great importance for the field. For more information on the datasets released in previous years, please see http://usewod.org/data-sets.html.

Special thanks got to Open Link Software for providing us with DBpedia logs as well as Ruben Verborgh for access to Linked Data Fragments usage data.

The data is available via the University of Southampton ePrints repository and has been assigned a DOI to allow for consistent referencing in publications.

PROGRAM COMMITTEE

- Elena Demidova
- Stefan Dietze
- Christoph Trattner
- Geert-Jan Houben
- Eero Hyvnen
- Irini Fundulaki

- Agnieszka Lawrynowicz
- Pasquale De Meo
- Johan Oomen
- Tom De Nies
- Efstratios Kontopoulos
- Arjen de Vries
- Agis Papantoniou
- Fabian Floeck

WORKSHOP ORGANIZERS

- Bettina Berendt - KU Leuven, Belgium
- Laura Hollink - VU Amsterdam, The Netherlands
- Markus Luczak-Roesch - University of Southampton, United Kingdom